Taking Stock of Survey-based Measures of Bureaucratic Red Tape: Mere Perceptions or More than Meets the Eye?

Sanjay K. Pandey, Rutgers University–Newark, skpandey@andromeda.rutgers.edu
&
Justin Marlowe, University of Washington, jmarlowe@washington.edu

Paper presented at the 10th Public Management Research Conference, Columbus, Ohio, October 1-3, 2009

Abstract
Despite an upsurge of research on bureaucratic red tape, a central issue about conceptualization and measurement of bureaucratic red tape remains unresolved. The most widely used measure of red tape is based on key informant reports about bureaucratic red tape in the organization. The starkest objection to such measures is that key informant reports are mere perceptions, subject to a variety of influences, and therefore cannot be accepted as reliable indicators of organizational reality. We assess the validity of key informant perception-based measures of bureaucratic red tape by employing the "anchoring vignettes" technique developed by King, et. al (2004). Anchoring vignettes are unique because they permit comparison of general perceptions with concrete reality-based scenarios. This method allows us to determine the level of "differential item functioning" (DIF) – or the degree to which respondents perceive the survey question differently - on survey questions that measure red tape. Although we find evidence for differential item functioning in how respondents perceive and report on red tape, it does not have a material effect.

Introduction

As research on bureaucratic red tape has burgeoned so have critiques (e.g. Kelman, 2008; Luton, 2007) and caveats about self-reported perceptual measures in red tape scholarship (e.g. Walker and Brewer 2008, 2009). The critiques take issue with the kind of judgments (multiple and complex) survey respondents are asked to make in order to answer questions on bureaucratic red tape and as a result point to the sheer unreliability of self-reported measures[1]. Caveats by

[1] Survey research is often subjected to "broad-brush" critiques that are not necessarily theoretically well-developed. Spector's (2006: 222) comments in this regard are worth noting, "Interestingly, the concern for CMV seems to be raised almost exclusively when cross-sectional, self-report surveys are used. Monomethod studies (those using the same method for assessing all variables) using other approaches, such as reports about other people (e.g., assessment centers or job performance ratings), are less criticized for the same shortcoming, although some have noted that source bias can be a problem in these other domains … This automatic criticism of the cross-sectional self-report has become invoked so broadly and often so automatically that I argue it has achieved the status of a methodological urban legend."

scholars studying red tape, unsure of where the critical thrust and parry may lead, dutifully note the use of perceptual measures as a limitation. Walker and Brewer (2009: 442), for example, note, "Our data are limited in another way, providing mainly perceptual measures..." Although there are isolated attempts to emphasize the value of perceptual measures in public management research (e.g. Andrews 2009; Brewer 2006; Pandey and Welch 2005; Yang and Pandey 2009) by linking the use of perception-based measures to established theoretical frameworks in organizational studies that emphasize the enacted (or socially constructed) aspects of organizational reality (Boyd, Dess, and Rasheed, 1993; Hambrick 1982; Simon 1953; Weick 1995), it is clear that there is some unease about the use of perception-based measures of red tape.

The reservations about relying on perceptions to assess red tape are by no means of recent vintage. Noted public administration scholars over time have made the point that assessments of bureaucratic red tape pit one person's opinion against another's (e.g. Kaufman 1977; Waldo 1946). Gouldner (1952), the renowned sociologist, goes farther than others in this matter and dismisses the possibility of coming up with an inter-subjectively valid scheme for measuring red tape. According to Gouldner, beneath the patina of concern with efficiency and burdensome administrative procedures, there is a fierce contest among competing social values. In other words, perceptions of red tape are not only about violation of the efficiency norm but may also be contaminated with respondent perspectives on other values such as equity, due process etc.[2]

---

[2] To be sure, it is possible to read more than violation of the efficiency norm in recent definitions of red tape (see Bozeman 1993, 2000; Pandey and Scott 2002). This is because Bozeman uses the phrase "no organizational or social benefit" in defining red tape, which may be seen as more encompassing than inefficient application of rules. A fuller and closer review of conceptual and operational definitions, however, makes clear the emphasis on inefficient application of rules is a central attribute of definitions of red tape.

This skepticism about the possibility of devising meaningful measures of red tape has not deterred scholars and questionnaire-based measures of perceptions of red tape have been extensively used since the 1990s (Baldwin 1990; Bozeman, Reed, and Scott 1992; Coursey and Pandey 2007; Rainey, Pandey, and Bozeman, 1995; Welch and Pandey 2007). The foundational work and arguments advanced by Bozeman (Bozeman 1993; 2000) in this regard have had considerable influence. Bozeman (1993) argues that it is possible to study burdensome rules and procedures that confer no organizational (or social) benefit. Despite some struggle in coming up with ways to distinguish burdensome from necessary (e.g. Pandey and Bretschneider, 1997), scholarship on bureaucratic red tape has pressed ahead. This scholarship makes the implicit assumption that we can indeed rely on managerial perceptions to assess bureaucratic red tape. Given the right question wording, the expectation is that managerial perceptions incorporate judgments about violation of the efficiency norm more than about other concerns.

How does the recently resurgent scholarship on bureaucratic red tape that employs perception-based measures of red tape address the issues raised by earlier scholars such as Gouldner, Kaufman, Waldo among others? The short answer is that it does not do so directly. The closest head-on engagement of this issue – to the best of our knowledge -- is by Pandey (1995: 5-8) who argues that implicit in classical critiques is a "straw dog" that "by pointing the divergence between clients' and managers' conception of red tape" concludes that it is not possible to get at inter-subjectively valid reality of red tape. Pandey recommends instead that red tape be examined as socially constructed reality that is dependent on the individual's social role. In other words, we should study red tape that managers face and also red tape that clients (or other stakeholders) face rather than continue the quixotic quest to label something red tape if and only if managers and other stakeholders can come to full agreement.

Indeed, public management scholars who are interested in managerial behavior have focused primarily on managerial perceptions of red tape, specifically in the domain of public personnel systems. Although this is a seemingly reasonable resolution, the resultant approach relying on survey questions to measure personnel red tape potentially falls short on two counts: 1) There are no *a priori* grounds for believing that individuals do not employ values other than administrative efficiency to answer these questions; and 2) Individuals are able to competently make the kinds of judgments necessary to answer these questions. If individuals rely on other values in answering questions about red tape – as Gouldner (1952) suggests – then we should see variation that is not rooted in the administrative burdens that are at play. Also, if individuals cannot be expected to make accurate judgments then again we should see variation that is not related to the nature of administrative barrier faced. Although survey questions continue to be used widely and a voluminous and impressive literature is developing around red tape (e.g., Baldwin, 1990; Feeney and Bozeman, 2009; Brewer and Walker, 2009; Pandey and Moynihan, 2006; Rainey, Pandey, and Bozeman 1995; Walker and Brewer 2008, 2009; Welch and Pandey 2007), it is hard to know whether or not (or more appropriately the extent to which) we can repose trust in it.

In order to address this question, we rely on the anchoring vignette technique developed by King et al. (2004). Vignettes, as the name suggests, present short real world-based scenarios to the respondents. They have been used in the past in survey research for a number of purposes. For example, Martin and Polivka (1995) report on the use of vignettes to modify and redesign the Current Population Survey (CPS). As Tourangeau, Rips, and Rasinski (2000: 324) note the "vignettes revealed that respondents' definitions of work sometimes differ markedly from the CPS definition." We use the King et al. anchoring vignette technique to test whether respondent

assessment of bureaucratic red tape, in fact, reflect efficiency and other values with the way red tape concept is defined in the public management literature (for overviews, see Bozeman 1993; Pandey and Scott 2002). Specifically, we develop three anchoring vignettes that represent three different levels of red tape in promotion decisions. Using these vignettes, we assess and correct a red tape questionnaire item, measuring red tape in promotion decisions, for differential item functioning (DIF). If the implicit claims of recent scholarship on bureaucratic red tape, employing survey questions, about the validity of survey question-based approach are true then we expect the following: 1) respondents will be able to order vignettes accurately; and 2) More importantly, results based on DIF-corrected scores would not be different from results based on raw and uncorrected scores.

Rest of the paper is structured as follows. First, we provide a brief overview of measurement of personnel red tape in public administration literature and the specific approach we take in this study. Next, we describe the data collection process and present the results of our analyses. We conclude the paper by discussing key findings, implications, and suggestions for future research.

Measuring Personnel Red Tape through Survey Questions

Although bureaucratic red tape can manifest itself in different organizational systems (Bozeman 2000; Coursey and Pandey, 2007; Pandey and Scott 2002), personnel (or human resources) red tape has attracted the most research attention (e.g., Baldwin 1990; Bozeman, Reed, & Scott, 1992; Bretschneider 1990; Feeney and Rainey, forthcoming; Pandey and Kingsley 2000; Pandey and Moynihan 2006; Walker and Brewer 2009). Survey questions typically focus on common tasks in personnel administration such as recruitment, discipline and termination, administration of pay raises, and promotions. A number of measurement strategies

have been tried in questions such as asking respondents to estimate the amount of time needed to complete tasks, number of decision-makers who need to sign-off for a task to be completed, and also measures that are almost solely based on respondent's attitudes (see Rainey, Pandey, & Bozeman, 1995 for the use of full range of these measures).

Over time, questionnaire items based solely on respondent attitudes about key personnel tasks have come to be used more often than alternative measures. Typically, respondents are asked to rate a number of personnel tasks on a Likert scale which are then used either as a summative scale (e.g. Pandey and Kingsley, 2000; Pandey, Coursey, and Moynihan 2007) or individually with a focus on a specific aspect of the personnel process (e.g. Feeney and Rainey, forthcoming; Walker and Brewer 2009). Rainey (1979) devised the oldest and most widely used set of five questionnaire items to measure personnel red tape. Recently, Wright and Pandey (forthcoming) use a slightly modified version of these items with the modification changing the focus of questions from managerial employees to all employees.

To assess whether or not attitudinal items capture what researchers want to capture about organizational reality, studying a single item can be as instructive as the full set. Therefore, we craft anchoring vignettes to go along with extent to which respondents agreed or disagreed with the following statement, "Personnel rules on promotion make it hard for a good employee to move up faster than a poor one." In answering this question, respondents are asked to make a clear value judgment that faults rules for the inability of good employee to obtain career advancement. The research tradition on bureaucratic red tape that follows Bozeman's lead (Bozeman 1993; 2000) would argue that when respondents answer this question, they are reporting on their organizational experience with "inefficient rules" pertaining to the promotion process.

Earlier theorists, however, may perhaps posit that in answering this attitudinal question, the respondents not only rely on an efficiency norm but may also bring in other values and propensities. Merton (1940), for example, posits the "bureaucratic personality" to describe a tendency to value rules and procedures over outcomes. The hallmark of bureaucratic personality is the acceptance of following rules as an important value. In Merton's words, "Discipline, readily interpreted as conformance with regulations, whatever the situation, is not seen as a measure designed for specific purposes but becomes an immediate *value* in the life-organization of the bureaucrat (1952: 355-56, emphasis added)". We would expect "inveterate" rule-followers, therefore, to be unable to notice a violation of the efficiency norm. Or at the very least have their assessment considerably dampened or tempered by their reverence for rules.

A counter-model is provided by Victor Thompson (1961). Thompson proposes a personality type he labels "bureautic". The bureautic, according to Thompson (1961: 176), "tends to lose sight of the organization as an instrument for accomplishing goals, as a structure of instrumental functions and relationships." Where Merton's bureaucratic personality finds rules to be invisible, Thompson's bureautic personality does not see the value of organizational rules and as such is pre-disposed to rate rules as burdensome. Although others (e.g. Dimock, 1952; Selznick 1949) have proposed models that share similar goals, the recent surge of red tape scholarship in the Bozeman tradition has tended to dismiss issues raised by these scholars. Pandey and Kingsley (2000: 783), for example, provide the following assessment of Thompson's bureautic, "[The] portrait of the bureautic individual stretches the bound of credulity and is too idiosyncratic… The complaint of red tape is raised too frequently to be ascribable to a somewhat whimsical and potentially rare type of personality."

This assessment, however, is a considered judgment and is not based in an empirical testing of how individuals answer survey questions. To our knowledge, there have been no attempts to examine if there is systematic over-rating or under-rating of red tape by individuals. The closest such effort we can think of is by Bozeman and Rainey (1998) who report that some individuals prefer more rules than others. They do not, however, try to examine if this translates into how such individuals answer survey questions on red tape and to explore corrections. We take these steps in this paper.

Data Collection

The data for this study were collected in Phase 4 of the National Administrative Studies Project (NASP-IV) using a survey administered to a nationwide sample. The theoretical population of interest for NASP-IV was comprised of senior managers (both general and functional) in US local government jurisdictions with populations over 50,000. The general managers included the city manager and assistant/deputy city managers. Functional managers included in the study headed key departments, namely, Finance/Budgeting, Public Works, Personnel/HR, Economic Development, Parks and Recreation, Planning, and Community Development.

The sample design and construction for the NASP-IV study was aided by the International City/County Management Association (ICMA). Based on the study criteria, ICMA compiled a list of potential respondents and the NASP-IV team used publicly available information to verify each respondent and identify a working email address. These efforts resulted in 3,316 individuals in the study sample. Each respondent in the study sample received an initial letter through US Mail which introduced the study and was directed the respondent to

complete the survey available on the study website using an assigned participation code. After the initial letter via US Mail, multiple methods were used in follow-up efforts to contact the respondents – e-mail, fax, and phone calls.

When the study concluded 1,538 of the 3,316 had responded, for a response rate of 46.4%. The anchoring vignettes (described below) were administered to a fraction of respondents who were randomly chosen to answer the anchoring vignette module in the survey. Pushing the vignettes to a portion of the sample rather than the entire sample is well-established in other applications of the anchoring vignettes technique. The advantage of doing so is efficiency. Vignettes consume space in the survey instrument and increase the time required to complete the instrument.[3] Following the convention in the anchoring vignettes literature, we also randomized the order in which anchoring vignettes were presented to the respondents. Of 1,538 respondents, 307 completed the vignettes. These 307 observations are used in this analysis.

Assessing Red Tape Perception in Promotion with Anchoring Vignettes

Anchoring vignettes are a tool to address the persistent problem in survey research of differential item functioning (DIF). DIF has been characterized several ways, but a simple definition is that DIF happens when respondents understand the same survey question in different ways. This is an acute problem in cross-cultural survey research on abstract ideas like fairness, efficacy, justice, etc. in different ways (King et al. 2004) because people in different societies experience these concepts in different ways. A "just" outcome in one society, for example, might be decidedly "unjust" in another, and when faced with the same question about their respective justice systems respondents in these different cultures might supply different

[3] Formal assessments of this technique show that vignette responses from a randomly selected sub-sample of respondents are sufficiently similar to those from the full sample (King, et. al 2004).

responses. This results in non-random measurement error that cannot be corrected by simply writing clearer survey questions.

In this analysis we equate DIF with the problem of multiple values embedded in red tape perceptions. Cross-cultural comparisons of perceptions are problematic because different cultures have different implicit norms and values that shape their members' world view. In our view, this same problem might be present in perceptions of red tape. Respondents with certain backgrounds or working within certain organizational cultures might view a particular personnel rule as an onerous bureaucratic procedure that causes inefficiencies, where others with different experiences might view it without concern for efficiency, and/or with concern for values other than efficiency like due process, fairness, and equity.

Anchoring vignettes address this problem by allowing us to strip away the efficiency dimension of HR red tape perceptions. This is accomplished by fixing those perceptions of efficiency to a common experience, thus reducing or eliminating variation due to divergent ideas about what efficiency means. Whatever DIF remains in this "DIF-corrected" HR red tape measure is, we believe, due to other values acting on red tape perceptions.

The key implication, and the thrust of our analysis, is that if the determinants of variation in the DIF-corrected measure are similar to those of the uncorrected measure, then HR red tape perceptions are free of the bias due to different ideas about how human resources red tape affects organizational efficiency. A corollary is that if the DIF-corrected measure performs similar to the non-corrected measure, then the influence of other values on red tape perceptions is probably more random than systematic. This would imply that concern about bias due to systematic divergence along dimensions other than efficiency is also less problematic than some might believe.

These corrections are accomplished by presenting survey respondents with short stories – "vignettes" – that place a perception-based question in a specific context. Comparing respondents' answers on the vignettes to own answer helps to correct DIF attributable to different experiences and perceptions with that particular dimension of the question. Put differently, we make exactly clear what "efficiency" means, and we compare the survey responses to that baseline. Vignettes can cover as many dimensions of the concept as the researcher would like and as the survey instrument will allow. We focus on only the efficiency dimension because, as discussed above, some of the key critiques of contemporary red tape measure attribute bias in red tape perceptions to varying perceptions about how red tape affects organizational efficiency.

The most important part of our empirical analysis is a comparison of results from two multivariate analyses of the determinants of red tape – one that uses the standard personnel red tape perception measure as its dependent variable, and another that uses the DIF-corrected measure as its dependent variable. If the results of these two analyses are not similar, then the raw personnel red tape measure is likely affected by DIF attributable to variation along dimensions other than efficiency. By contrast, similar results across the two sets of estimates would imply that DIF due to values other than efficiency has no systematic effect on the inferences we can appropriately drawn from analysis of individual perception-based red tape measures.

With these considerations in mind, we employed the following analytical strategy. We first asked respondents the standard red tape perception measure with respect to their organization: "Personnel rules on promotion make it hard for a good employee to move up faster than a poor one." Answers followed a five point scale where 1 = "strongly disagree" and 5 =

"strongly agree." Respondents were then presented three vignettes that briefly described different hypothetical promotion scenarios and were asked to assess the level of human resources red tape experienced by the managers in each vignette. That assessment followed the same wording and scale as the personnel red tape perception measure. Those vignettes were as follows:

- (1) "Gene would like to promote a highly effective young employee. The city's personnel rules provide managers broad discretion in hiring and promotion. Within one month Gene requests a promotion for the employee, that request is approved, and the new employee begins in the new position."

- (2) "Chris would like to promote an excellent employee. Chris is advised to nominate a more experienced employee because the city's personnel system rewards tenure. The excellent employee is nominated, the promotion process takes 9 months, but the promotion is approved. Chris is advised to 'play by the rules' next time."

- (3) "Terry would like to promote an excellent, but relatively new employee. Personnel rules require that experienced employees should be promoted first, but those rules allow for exceptions. Terry requests such an exception, and that request is denied. Terry repeats the request six months later and is denied again. Terry gives up on the idea of rewarding the excellent employee through early promotion."

These vignettes were carefully crafted to make certain they speak only to the efficiency dimension. To do this we limited the details of each to two main considerations - the time the promotion process took, and the latitude each manager was afforded to promote their chosen employee. The vignettes were reviewed by a panel of public management scholars. We also pre-tested the vignettes with a pilot group of survey respondents and adjusted the question wording several times according to their input.

Proper application of the anchoring vignettes technique requires some assumptions about the implicit ordering of the vignettes relative to the assessment question. For these vignettes the assumed ordering was as presented above: "Gene" experienced the least red tape, followed by "Chris" and "Terry." Gene experiences comparatively little red tape, and in Terry's case the red tape is so severe that s/he abandons his/her effort to promote the desired candidate. As discussed

below, responses were generally consistent with this assumed ordering. This suggests the vignettes fairly characterize the variation along the efficiency dimension.

Respondents' red tape self-assessment answers were then recoded relative to their answers to the vignettes to produce a "DIF-corrected" red tape measure. This new, DIF-corrected measure has seven categories. If a respondent answered the vignettes according to our assumed order of Gene-Chris-Terry, and their answer to the self-assessment question was less than Gene, that respondent would receive a DIF-corrected red tape measure of 1, 2 if equal to Gene, 3 if more than Gene but less than Chris, and so forth. Figure 1 presents our respondents' vignette orderings. For consistency, vignettes are identified in terms of their assumed ranking from least to most red tape: Gene = 1, Chris = 2 and Terry = 3. "Ties," or the same response for two or more vignettes, are identified in brackets.

[FIGURE 1 ABOUT HERE]

A majority of the respondents (175 out of 307, or 57%) identified vignette 1 (Gene) as having the least red tape, 60 gave both correct and distinct (i.e. no ties) responses, and 235 (76%) identified the correct ordering if we assume that equivalence between 2 and 3 is not an "incorrect" response. These results suggest the respondent orderings were generally consistent with our assumed orderings, and that the vignettes provide defensible anchors for DIF-correction.

Many respondents gave identical answers for vignettes 1, 2, and/or 3. In the parlance of anchoring vignettes these are "ties." Ties are difficult to use when combining the self-assessment with the vignettes because they span multiple categories. That is, if a respondent gave the same answer for the Gene and Chris vignettes, that response could be re-coded between 1 and 4,

depending on how we choose to "break" that tie.[4] The same applies to incorrect orderings, or responses that deviated from our assumed "1,2,3" ordering. Incorrect answers account for 23% of our responses. Ties and incorrect orderings are not uncommon in applications of anchoring vignettes.[5] Moreover, our use of the parametric estimator described in the next section allows us to extract additional relevant information from those ties.

Multivariate Model and Results: Comparison of Raw and DIF-corrected Scores

Our core analytical strategy was to compare the results of a multivariate analysis of the determinants of red tape perceptions using both the traditional red tape perception measure to those same results using a DIF-corrected measure. Large changes in the size, direction, and/or statistical significance of the coefficients across the two models would suggest that DIF due to divergent ideas about the red tape-efficiency relationship is a potential source of bias in red tape measurement.

To do this we first developed an empirical model of the determinants of red tape perceptions. That model was grounded in the extant red tape literature, and we designed to more or less replicate previous findings in this area. It includes the respondent's perceptions of the

[4] Several approaches have been suggested to correct this problem. One is to simply ignore the ties. The logic here is that if a respondent cannot provide a clear ordering of the vignettes, then recoding their self-assessment relative to the vignette adds little information relevant to the DIF correction. A second approach is to allocate equal portions of each case to the DIF-corrected variable categories spanned by the tie. For example, if a respondent answered "Agree" to the self-assessment question and to both the Chris and Terry vignettes, then half of their response would be allocated to category 4 and half to category 6. In the presence of many ties, as we have here, equal allocation produces an essentially uniform distribution of the DIF-corrected measure. This adds little new insight. A third approach developed by King and Wand (2007) is known as censored ordered probit allocation. In short, this method distributes each tied response according to the proportion of respondents who gave the same answer on the self-assessment question relative to the same discrete categories for the vignettes that cover the categories spanned by the tied response. The intuition here is that respondents with the same self-assessment and similar, albeit tied vignette responses have more in common than respondents with the same self-assessment but dissimilar vignette responses. This method is preferable to the uniform allocation method because it extracts additional information from the ties.

[5] For a recent treatment of techniques to reduce the incidence of ties see Hopkins and King (2009). Within the anchoring vignettes literature there is also an extensive array of tests for "vignette equivalence," or the idea that vignettes are substitutable. A separate set of procedures tests whether vignettes address single or multiple dimensions of a concept. For brevity, we do not report the iterations of those tests on our vignettes. See, for instance (King and Wand, 2007).

organization and several demographic variables believed to influence perceptions of red tape. For further discussion on these variables, their hypothesized effects, and the logic behind their inclusion, see (Pandey and Moynihan 2006; Pandey and Welch 2005; Pandey and Kingsley 2000). Descriptive statistics for the model variables are reported in Table 1.

[TABLE 1 ABOUT HERE]

The model includes measures of the respondent's perception of three aspects of their organization. *Developmental culture* is comprised of three items: "My department is a very dynamic and entrepreneurial place. People are willing to stick their necks out and take risks;" "The glue that holds my department together is a commitment to innovation and development. There is an emphasis on being best" and "My department emphasizes growth and acquiring new resources. Readiness to meet new challenges is important." *Political support* is comprised of two items – "Most elected officials trust the organization" and "Most elected officials believe that the organization is effective." The final perception-based item, *goal clarity*, is comprised of "This organization's mission is clear to almost everyone who works here"; "This organization has clearly defined goals"; and "It is easy to explain the goals of this organization to outsiders." All survey question items follow a five point scale where 1 = "strongly disagree" and 5 = "strongly agree." These concepts were measured using either previously validated measures or minor adaptations thereof (Gianakis and Wang 2000; Rainey, 1983; Zamutto and Krakower 1991). The model also includes *organization size*, measured as the total number of employees within the organization.

It also includes measures of individual job satisfaction and other individual characteristics. *Job satisfaction* is comprised of two items derived from the Michigan Organizational Assessment Questionnaire: "In general, I like working here" and "In general, I like the type of work I do." Age is the respondent's age in years. Education is a four item variable where 1 = "some college"; 2 = "bachelor's degree"; 3 = "master's degree"; and 4 = post-graduate degree. *Current tenure* is the number of years the respondent has been with the organization, *current position* is the number of years the respondent has been in their current position, and *salary* is the respondent's current annual salary, coded as 1 = " Less than $50,000," 2 = $50,000 to $75,000, 3 = $75,000 to $100,000, 4 = $100,000 to $150,000, and 5 = $150,000 or more. The model also includes dummy variable indicators for women, non-white respondents, and respondents identified as human resources professionals. This final variable was included to control for the fact that in most organizations human resources professionals create and administer human resources rules. As such, we expect they are less likely to perceive those rules as inhibiting efficiency.

We estimated the model of the basic red tape self-assessment measure with ordered probit. For the DIF-corrected measure we used conditional hierarchical ordered probit (or "chopit"; see King, et. al. 2004). The chopit routine is different from traditional ordered probit in that it allows the model cut-points to vary across respondents. Those cut-points are estimated with a pre-specified model of observed independent variables. We chose to model them as a function of the individual-specific variables in our model: age, education, race, and sex. Those cut points are then estimated relative to both the vignettes and the response to the self-assessment measure. As such, the chopit routine estimates separate coefficients for each independent variable across the thresholds for each cut-point. For ease of comparison we present only the

"betas" (King, et. al. 2004), or estimates of independent variables that associate with an assumed overall model intercept of 0. The two sets of estimates are presented in Table 2. The left column reports the Chopit model estimates, and the right column reports the estimates from a traditional ordered probit model.

The results presented in Table 2 show that the estimates for the two models are more or less the same. With two exceptions, all statistically significant coefficients in the ordered probit model remain significant and have the same basic magnitude and direction in the Chopit model. One exception is current tenure, which becomes significant in the Chopit model, and the second is organization size, which becomes non-significant in the Chopit model. But in both cases the coefficients remain marginally significant, and the effect on their size is negligible.

It is important to note that the DIF-corrected dependent variable does, in fact, result in more precise estimates. Standard errors for the coefficients in the Chopit model are lower overall, and in some cases the gain in precision is appreciable. That said, even with those gains in precision, we reach the same basic conclusions about the factors that shape red tape perceptions. For most researchers, we suggest, the cost of collecting the additional data necessary to implement the DIF corrections will likely outweigh the benefits of the additional precision of estimation from those corrections.

Conclusion

As theoretically ambitious research on the red tape concept has continued to cumulate, the reliance on perceptions in survey questions to measure red tape has hung as a big question mark. Although some (e.g. Spector 2006) have questioned the value of broad-brush critique of survey research, some of the critique of perception-based measures of red tape is rooted in prior theoretical work which argues that in answering such questions respondents may deploy a range

of values in addition to potential violation of the efficiency norm. Our analysis, using the anchoring vignette technique, shows that survey questions measuring red tape are viewed as intended by researchers and that respondents do not necessarily invoke other values in answering these questions. While perception-based measures of red tape may suffer from other limitations, it seems that respondent reliance on other values is not one of them.

One limitation of our research is that our vignettes only capture the time and the decision-making latitude aspects of red tape. Other vignettes could capture other aspects, but would require more time and would come at a higher cost. More vignettes may get more precision, but at the risk of inadvertently measuring other dimensions of the concept in question. We are also limited by the fact that the study focuses one aspect of personnel red tape. One might obtain different conclusions if other kinds of red tape, or even different aspects of personnel red tape, were studied. (e.g., the time between the decisions to promote until promotion actually happens, the degree to which red tape impairs organizational effectiveness).

This research suggests future avenues as well some that relate to other potential applications of the anchoring vignette technique in public administration literature and others about advancing the anchoring vignette methodology. On a methodological front, it would be helpful to understand the extent to which DIF is shaped by factors at different levels such as role-specific, demographic-specific, or organization-specific. This technique can also be used to study many other perception-based issues and concepts in public administration literature such as public service motivation, goal and role clarity.

References

Andrews, Rhys. 2009. Organizational Task Environments and Performance: An Empirical Analysis. *International Public Management Journal*, 12(1): 1-23.

Baldwin, J. Norman. 1990. Perceptions of Public vs. Private Sector Personnel and Informal Red Tape: Their Impact on Motivation. *American Review of Public Administration*. 20: 7-28.

Boyd, Brian, Gregory Dess, and Abdul Rasheed. 1993. Divergence between archival and perceptual measures of the environment: Causes and consequences. *Academy of Management Review* 18: 204–26.

Bozeman, Barry. 2000. *Bureaucracy and Red tape*. Upper Saddle River, NJ: Prentice Hall.

Bozeman, Barry, 1993. A theory of government red tape. *Journal of Public Administration Research and Theory*. 3: 273-303.

Bozeman, Barry, and Hal G. Rainey. 1998. Organizational rules and bureaucratic personality. *American Journal of Political Science*, 42(1), 163-189.

Bozeman, Barry, and Patrick Scott. 1996. Bureaucratic red tape and formalization: Untangling conceptual knots. *American Review of Public Administration*, 26(1): 1-17.

Bozeman, Barry, Pamela Reed, and Patrick Scott. 1992. Red tape and task delays in public and private organizations. *Administration & Society*. 24(3): 290-322.

Bretschneider, Stuart I. Management Information Systems in Public and Private Organizations: An Empirical Test. *Public Administration Review*. 50: 536-545.

Brewer, Gene A. 2006. All Measures of Performance are Subjective: More Evidence on U.S. Federal Agencies. In *Public Service Performance*, ed. G. A. Boyne, K. J. Meier, L. J. O'Toole and R. M. Walker. Cambridge, England: Cambridge University Press.

Brewer, Gene A. and Richard Walker. Forthcoming. The Impact of Red Tape on Governmental Performance: An Empirical Analysis. *Journal of Public Administration Research and Theory*.

Coursey, David H., and Sanjay K. Pandey. 2007. Content Domain, Measurement, and Validity of the Red Tape Concept: A Second-Order Confirmatory Factor Analysis. *American Review of Public Administration*. 37: 342-361.

Dimock, Marshall. 1952. Bureaucracy self-explained. In *Reader in Bureaucracy* (Eds.) R.K. Merton, A.P. Gray, B. Hockey, and H.C. Selvin. Glencoe, IL: Free Press, pp. 397-406.

Feeney, Mary K. and Hal G. Rainey. Forthcoming. "Personnel Flexibility and Red Tape in Public and Nonprofit Organizations: Distinctions due to Institutional and Political Accountability." *Journal of Public Administration Research and Theory*.

Feeney, Mary K. and Barry Bozeman. 2009. Stakeholder Red Tape: Comparing Perceptions of Public Managers and Their Private Consultants. *Public Administration Review*. 69(4): 710-726.

Gouldner, Alvin. 1952. Red tape as a social problem. In *Reader in Bureaucracy* (Eds.) R.K. Merton, A.P. Gray, B. Hockey, and H.C. Selvin. Glencoe, IL: Free Press, pp. 410-418.

Gianakis, Gerasimos., and Xiaohu Wang 2000. Decentralization of the purchasing function in municipal government: A national survey. *Journal of Public Budgeting, Accounting and Financial Management*, 12(3), 421-440.

Hambrick, Donald. 1982. Environmental scanning and organizational strategy. *Strategic Management Journal*. 3 :159–74.

Hopkins, Daniel and King, Gary Forthcoming. Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability. *Public Opinion Quarterly*.

Kaufman, Herbert. *Red Tape: Its Origins, Uses, and Abuses*. Washington, DC: Brookings.

Kelman, Steven. 2008. The 'Kennedy School School' of Research on Innovation in Government. In *Innovations in Government: Research, Recognition, and Replication*. Ed. Sandford Borins. Washington, DC: Brookings Press, pp. 28-52.

King, Gary, Christopher J.L. Murray, Joshua A. Salomon, and Ajay Tandon 2004. Enhancing the Validity of Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review* 98(1): 191-207.

King, Gary and Wand, Jonathan 2007. Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes. *Political Analysis* 15: 46-66.

Luton, Larry S. 2007. Deconstructing Public Administration Empiricism. *Administration & Society*, 39(4): 527 - 544.

Merton, Robert K. 1952. Bureaucratic structure and personality. *Social* In *Reader in Bureaucracy* (Eds.) R.K. Merton, A.P. Gray, B. Hockey, and H.C. Selvin. Glencoe, IL: Free Press, pp. 361-371.

Pandey, Sanjay K., David H. Coursey, and Donald P. Moynihan. 2007. Organizational effectiveness and bureaucratic red tape: A multi-method study. *Public Performance and Management Review* 30 (3): 398–425.

Pandey, Sanjay K., and Donald P. Moynihan. 2006. Bureaucratic red tape and organizational performance: Testing the moderating role of culture and political support. In *Public Service Performance*, ed. G. A. Boyne, K. J. Meier, L. J. O'Toole and R. M. Walker. Cambridge, England: Cambridge University Press.

Pandey, Sanjay K. 1995. *Managerial Perceptions of Red Tape* Syracuse University: Unpublished PhD Dissertation.

Pandey, Sanjay K., and Eric W. Welch. 2005. Beyond stereotypes: A multistage model of managerial perceptions of red tape. *Administration & Society* 37(5): 542-575.

Pandey, Sanjay K., and Patrick G. Scott. 2002. Red tape: A review and assessment of concepts and measures. *Journal of Public Administration Research and Theory* 12 (4):553-580.

Pandey, Sanjay K., and Gordon A. Kingsley. 2000. Examining red tape in public private organizations: Alternative explanations from a social psychological model. *Journal of Public Administration Research and Theory*, 10(4), 779-799.

Pandey, Sanjay K., and Stuart I. Bretschneider. 1997. The impact of red tape's administrative delay on public organizations' interest in new information technologies. *Journal of Public Administration Research and Theory* 7:113–30.

Rainey, Hal G. 1983. Public agencies and private firms. *Administration and Society* 15:207–42.

Rainey, Hal G. 1979. Perceptions of incentives in business and government: Implications for civil service reform. *Public Administration Review*. 39: 440-448.
Martin, Elizabeth and Anne E. Polivka. 1995. Diagnostics for redesigning survey questionnaires: Measuring work in the Current Population Survey. *Public Opinion Quarterly*. 59: 547-567.

Rainey, Hal G., Sanjay K. Pandey, and Barry Bozeman. 1995. Research note - Public and private managers perceptions of red tape. *Public Administration Review* 55 (6):567-574.

Selznick, Philip. 1949. *TVA and the Grassroots*. Berkeley, CA: University of California Press.

Simon, Herbert A. 1953. Birth of an organization: The economic cooperation administration. *Public Administration Review*. 13:227–36.

Spector, Paul E. 2006. Method Variance in Organizational Research: Truth or Urban Legend? *Organizational Research Methods*. 9(2): 221-232.

Thompson, Victor (1961). *Modern organization*. New York: Knopf.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response.* New York, NY: Cambridge University Press.

Waldo, Dwight. 1964. Government by procedure. In *Elements of Public Administration*. F.M. Marx (Ed.) Englewood Cliffs, NJ: Prentice Hall, pp. 381-399.

Walker, Richard M. and Gene A. Brewer. 2009. Can Management Strategy Minimize the Impact of Red Tape on Organizational Performance? *Administration & Society*, 41: 423 - 448.

Walker, Richard M. and Gene A. Brewer. 2008. An Organizational Echelon Analysis of the Determinants of Red Tape in Public Organizations. *Public Administration Review* 68(6): 1112 - 1127.

Weick, Karl. 1995. *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

Welch, Eric W., and Sanjay K. Pandey. 2007. E-Government and bureaucracy: Toward a better understanding of intranet implementation and its effect on red tape. *Journal of Public Administration Research and Theory* 17:379-404.

Wright, Bradley E. and Sanjay K. Pandey. Forthcoming. Transformational Leadership in the Public Sector: Does Structure Matter? *Journal of Public Administration Research and Theory*.

Yang, Kaifeng, and Sanjay K. Pandey. 2009. How Do Perceived Political Environment and Administrative Reform Affect Employee Commitment? *Journal of Public Administration Research and Theory*. 19: 335-360.

Zammuto, R. F., & Krakower, J. Y. 1991. Quantitative and qualitative studies of organizational culture. *Research in Organizational Change and Development*, 5, 83-114.

Figure 1: Distribution of Vignette Orderings

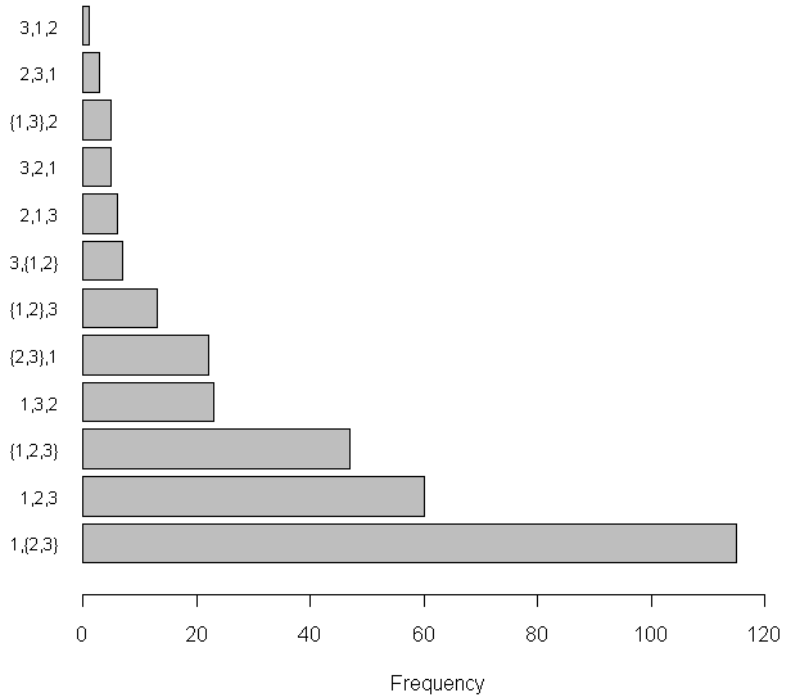

Table 1: Descriptive Statistics				
	Mean	**SD**	**Min**	**Max**
HR Red Tape	3.49	1.29	1	5
Job Satisfaction	9.20	1.25	2	10
Developmental Culture	11.45	2.55	4	15
Political Support	10.32	2.94	2	14
Goal Clarity	10.79	2.78	3	15
Organization Size	931.27	2.45	43	40014
Age	51.46	7.28	31	67
Education	2.92	0.85	1	4
Current Tenure	12.86	9.19	0	37
Current Position	7.35	6.63	0	33
Sex	0.29	-	0	1
Non-White	0.09	-	0	1
Salary	3.72	0.92	1	5

Table 2: Ordered Probit and Conditional Hierarchical Ordered Probit Estimates of Human Resources Red Tape Perceptions

	CHOPit	Ordered Probit
Job Satisfaction	-0.027	-0.038
	(0.055)	(0.091)
Developmental Culture	-0.051*	-0.085*
	(0.027)	(0.046)
Political Support	-0.036	-0.057
	(0.025)	(0.042)
Goal Clarity	-0.001	0.000
	(0.027)	(0.046)
Organization Size	0.129*	0.199
	(0.071)	(0.122)
Age	-0.006	0.001
	(0.012)	(0.016)
Education	-0.147	-0.156
	(0.091)	(0.123)
Current Tenure	0.018**	0.032**
	(0.008)	(0.014)
Current Position	-0.020	-0.033*
	(0.012)	(0.020)
Sex	-0.117	0.181
	(0.174)	(0.239)
Non-White	0.222	0.167
	(0.267)	(0.358)
Salary	-0.040	-0.062
	(0.070)	(0.121)
τ_1	-3.034	-3.460
τ_2	1.326	-1.942
τ_3	0.735	-1.331
τ_4	1.534	0.130
Log-Likelihood (df)	1699.739	903.49
Likelihood Ratio χ^2	83.92**	43.07**

N = 307
* = p< .1, ** = p< .05